Radiography 31 (2025) 103134

Contents lists available at ScienceDirect

Radiography

journal homepage: www.elsevier.com/locate/radi

Analysis of challenging mammographic cases demonstrates subtle reader group discrepancies

N. Clerkin [a,*], C. Ski [b], M. Suleiman [c], Z. Gandomkar [c], P. Brennan [c], R. Strudwick [a]

[a] University of Suffolk, Waterfront Building, 19 Neptune Quay, Ipswich IP4 1QJ, UK
[b] University of Sydney, Camperdown, NSW 2006, Australia
[c] School of Nursing and Midwifery, Queen's University Belfast, Belfast, UK

ARTICLE INFO

Article history:
Received 24 March 2025
Received in revised form
24 July 2025
Accepted 5 August 2025
Available online xxx

Keywords:
Mammography interpretation
Cancer detection
Advanced practice
Performance
Breast imaging

ABSTRACT

Introduction: High quality image interpretation is essential to detect early abnormalities on mammograms. A better understanding of the types of image characteristics that are most challenging to readers would support future education, as well as underpin advancements in AI modelling. This current work focuses on radiography advanced practitioners (RAP) to establish if RAPs and radiologists are challenged by the same characteristics.
Methods: This was a prospective, comparison study of radiographer and radiologist mammography readings. Using a cloud-based image interpretative platform and a 5 MP display, 16 radiographers and 24 radiologists read a test set of 60 mammograms with 20 confirmed cancer cases. Difficulty indices were calculated for each group based on error rates for each mammographic case. Unpaired Mann–Whitney tests were employed to compare error rates between various image characteristics. Spearman correlation analysis was used to establish if difficulty indices were associated with each cohort.
Results: Strong correlations for cancer and normal cases difficulty indices respectively ($r = 0.83$ CI:0.61–0.93) and ($r = 0.73$; CI:0.54–0.85) were shown between both groups. Greatest difficulty scores were shown for cases with soft tissue appearances as opposed calcifications ($p = 0.003$) and for cases without prior images, compared to those with ($p = 0.03$). No significant image characteristic differences were noted for the radiologists.
Conclusion: This early study acknowledges a strong correlation between radiologists and radiographers when determining which mammographic cases are difficult to interpret. However, radiographers appear to be more susceptible to varying cancer appearances as well as the non-availability of prior images with normal cases.
Implications for practice: The results should be helpful when tailoring educational strategies and developing augmented artificial intelligence (AI) solutions to support human readers.

Introduction

Early identification of breast cancer allows for timely interventions, potentially reducing aggressive treatments and mortality rates.[1] High quality mammography interpretation is essential to correctly identify abnormalities at an early stage, reducing the risk of false negatives or false positives that may lead to misdiagnosis or delayed diagnosis.[2] Understanding specifically which types of images are most challenging to mammographic readers can tailor and accelerate learning experiences thus ensuring maximum return from very limited available time during current pressurised clinical environments.[3] Understanding these difficult to interpret cases also supports the development and application of Artificial Intelligence (AI) models since AI modelling can be directed to what the human finds most challenging.[4,5]

Tailored education is not a new concept: learning customisation is an important component across many domains such as healthcare, education and sports training.[6–10] Better recognition of competencies already acquired, personal learning styles and specific needs of the individual have been proven to optimise and maintain performance. Such learning also encourages the individual to control their own learning process, promoting autonomy and engagement.[11] It is interesting that whilst medical imaging has advanced greatly in many ways, personalised learning remains limited. Nonetheless, in radiology and radiography some work has

* Corresponding author. 2 Everton Drive, Belfast, BT6 OLJ, UK.
E-mail address: noellefgeclerkin@gmail.com (N. Clerkin).

https://doi.org/10.1016/j.radi.2025.103134

already been achieved; this includes utilising technologies like AI and adaptive learning to create customised learning experiences.[12]

Recent clinical, academic and industrial partnerships have highlighted that with AI models, tailored mammographic test sets can be developed in an automated way, that can benefit breast image interpreters[13] with such endeavours relying on what makes things difficult for the reader. A recent three phased AI-based study, aimed to predict which mammographic cases were difficult for: experienced test set users engaging with new cases; inexperienced test set readers (with no history of test set reading) engaging with cases only read by other individuals and all readers engaging with cases that no one had ever engaged with.[12] This early work demonstrated much promise in being able to differentiate between cases that users would find difficult *before* they encountered the cases with mean area under the curve (AUC) values of 0.81 when differentiating between hard, versus easy to interpret cases.

This previous work however focussed on radiologist readings. To ensure that we fully understand which cases are challenging for*all* reader groups, work needs to acknowledge that in certain settings such as the United Kingdom (UK), enhanced, advanced and consultant radiographers who hold a qualification in image interpretation interpret mammograms. For the purpose of the study the term Radiography Advanced Practitioner (RAP) has been used to reflect this cohort of readers. As a first step towards this, the current study compares mammography readings of 16 RAPs and 24 radiologists to establish if the same cases present the same challenges for different reader groups. Acknowledging (if necessary) that difficult image presentations may be different for different reader types is a priority to optimise augmented AI development to meet user needs.

Objective

This paper therefore is responding to the current move towards personalised education and AI-based diagnosis. We cannot tailor educational strategies for human readers or reader cohorts nor can we direct AI-strategies optimally unless we better understand the types of medical images that represent most diagnostic difficulty. The current work therefore in the context of mammography, aims to examine two cohorts of breast readers – radiologists and RAPs, to firstly establish what type of images represent the biggest challenge and explore if this is consistent between these two reader groups.

Methods

Overview

This prospective study compared two cohorts of readers, radiologists and RAPs, each of whom read the same set of 60 cases, 20 of which demonstrated cancer, whilst the remaining 40 cases were normal. Readers were not informed of the ratio of positive and negative cases. All cases were verified to be typical mixed density screening cases by a senior radiologist responsible for training and quality assurance. The cases were considered by a senior reporting UK radiographer to ensure that cases represented a typical UK mix.

Ethics approval was granted from the lead researcher's affiliated university and consent was obtained from each reader through completion of an informed consent form.

Participants and recruitment

The study involved two groups of breast screen readers – radiologists (n = 24) and RAPs (n = 16). Recruitment was achieved through a number of avenues: direct contact with their workplace, social media platforms and international and national conferences. All readers reported for both symptomatic and breast screening cases and met minimum reading requirements in terms of qualifications and minimum annual reads in their respective countries. Radiologists were based in Australia, whilst RAPs were employed by the National Health Service (NHS) in the UK. This provided comparative professional perspectives and international practice variation. Reader details are shown in Table 1.

Mammographic cases

All readers read the same mammographic test set comprising of 60 cases gathered from the Australian Breast Screening Programme and displayed in random order. Each case included two standard cranial-caudal (CC) and two medial-lateral oblique (MLO) projections and all images demonstrated acceptable technical quality as verified by an independent consultant breast radiologist. Prior cases were also presented when available. Twenty of the cases contained a single biopsy-detected malignancy and the remaining 40 cases were normal or benign and verified with a normal screen two years later. Whilst the cancer cases consisted of a variety of invasive and in-situ types with appearances including asymmetric densities (n = 2), spiculated masses (n = 10) and indeterminate micro calcification positive for ductal carcinoma in situ (n = 8), for the purposes of this study we simply divided them into soft tissue masses (n = 12) and calcifications (n = 8). The 40 normal cases included benign findings such as duct ectasia, fibroadenomas, oil cysts and intra-mammary lymph nodes.

All cases were graded for mammographic density using Volpara density grading.

Details of these cases are shown in Table 2 (cancer) and 3 (normal) (see Table 3).

Image viewing

To simulate a clinical reporting environment and as recommended by the NHSBSP, participants undertook the study on a 5 MP reading workstation, in optimum lighting conditions suitable

Table 1
Reader demographics.

Reader type	Median value	Interquartile range
Radiologists		
Age	Not available	
How many years qualified	15	10–23
How many years reading mammographic images	10	10–18
Radiography advanced practitioners		
Age	45	35–55
How many years qualified	15	8–35
How many years reading mammographic images	8	30–69

Table 2
Details of the cases presenting with cancer.

Case	Cancer appearance (soft tissue mass or calcification)	Mammographic density BIRAD scale A(Fatty) - D (Extremely Dense)	Prior
Case 1	Calcification	D	Yes
Case 2	Soft tissue	A	Yes
Case 3	Soft tissue	B	Yes
Case 5	Soft tissue	B	No
Case 8	Soft tissue	B	Yes
Case 10	Soft tissue	C	Yes
Case 13	Soft tissue	B	Yes
Case 14	Soft tissue	D	Yes
Case 16	Soft tissue	B	No
Case 18	Calcification	B	Yes
Case 26	Calcification	B	No
Case 32	Calcification	B	No
Case 33	Calcification	C	No
Case 35	Soft tissue	A	Yes
Case 36	Calcification	B	Yes
Case 40	Soft tissue	B	Yes
Case 41	Soft tissue	C	No
Case 49	Soft tissue	B	Yes
Case 51	Soft tissue	C	Yes
Case 57	Calcification	A	Yes

Table 3
Details of the normal cases.

Case	Mammographic density BIRAD scale A(Fatty) - D (Extremely Dense)	Prior
Case 4	C	No
Case 6	B	Yes
Case 7	C	No
Case 9	B	No
Case 11	C	Yes
Case 12	B	Yes
Case 15	A	Yes
Case 17	B	Yes
Case 19	B	Yes
Case 20	C	Yes
Case 21	B	Yes
Case 22	A	No
Case 23	C	Yes
Case 24	B	Yes
Case 25	C	Yes
Case 27	B	No
Case 28	B	No
Case 29	B	Yes
Case 30	B	Yes
Case 31	C	Yes
Case 34	B	Yes
Case 37	A	Yes
Case 38	B	Yes
Case 39	C	No
Case 42	B	Yes
Case 43	D	Yes
Case 44	B	No
Case 45	B	Yes
Case 46	C	No
Case 47	B	Yes
Case 48	B	Yes
Case 50	A	Yes
Case 52	C	Yes
Case 53	C	Yes
Case 54	C	Yes
Case 55	D	Yes
Case 56	B	Yes
Case 58	C	Yes
Case 59	B	No
Case 60	C	Yes

for image interpretation. Images were accessed via the cloud using the DetectedX platform, an online elearning platform (DetectedX, Sydney Australia) in such a way that full native resolution was available at all times. The typical post-processing tools were available such as windowing, panning and magnification. Participants marked any area of an image that was indicative of cancer. Once marked, the reader rated their confidence using a one to five scale, with one representing complete confidence that the case was normal, two representing a benign case and five representing absolute confidence that a malignancy was present. Any marking located within a pre-set radius from a cancer appearance was considered correct, with radii set by expert radiologists with the relevant pathology reports. Readers could mark as often as they liked on any of the images. All readers were able to change any image judgements as often as they wished, until the stage when all cases were completed and all answers submitted. Information on the number of cases with an abnormality was not made available to the participant. The scoring system and image processing options available on the DetectedX platform was explained to each reader prior to commencing the study.

Data analysis

Each case was given a specific difficulty index as described by Rawashdeh et al.[14] Briefly, this was calculated as follows. For the cancer cases, the number of readers that did not correctly locate the cancer (gave a score of 1 or 2 using the scoring scale described above) was divided by the total number of readers. So, for example if 20 of the 24 radiologists did not identify the cancer for a particular case, the difficult index for that case would be $20/24 = 0.8$. The same process was applied to normal cases: for each case, the number of readers who incorrectly identified a normal case as having cancer (gave a score of 3–5) was divided by the total number of readers. This was done separately for the radiologist and radiographer cohorts.

Analyses were performed separately for the radiologist and radiographer cohorts and to test for data normality, the D'Agostino-Pearson omnibus test was employed. For cancer cases, difficulty indices were compared between different the two types of cancer appearance (soft tissue vs calcifications), level of breast density (A, B vs C, D) and the availability of prior images vs no prior images. For normal cases, density levels (A, B vs C, D) and the availability of prior images vs no prior images were analysed. Due to non-normal distribution of the data, the unpaired Mann–Whitney tests were employed for the above comparisons. Lastly, the Spearman correlation analysis was used to identify associations between radiologists and radiographer difficulty indices for cancer and normal cases separately.

Results

Cancer cases

The four cancer cases with the highest and lowest difficulty indices are shown respectively in Fig. 1a and b for radiologists and Fig. 2a and b for RAPs. Two cases overlapped for both cohorts (Cases, 5 and 35). Our findings demonstrate that for RAPs, significantly greater difficulty was found for those cancers with soft tissue appearances compared with those demonstrating

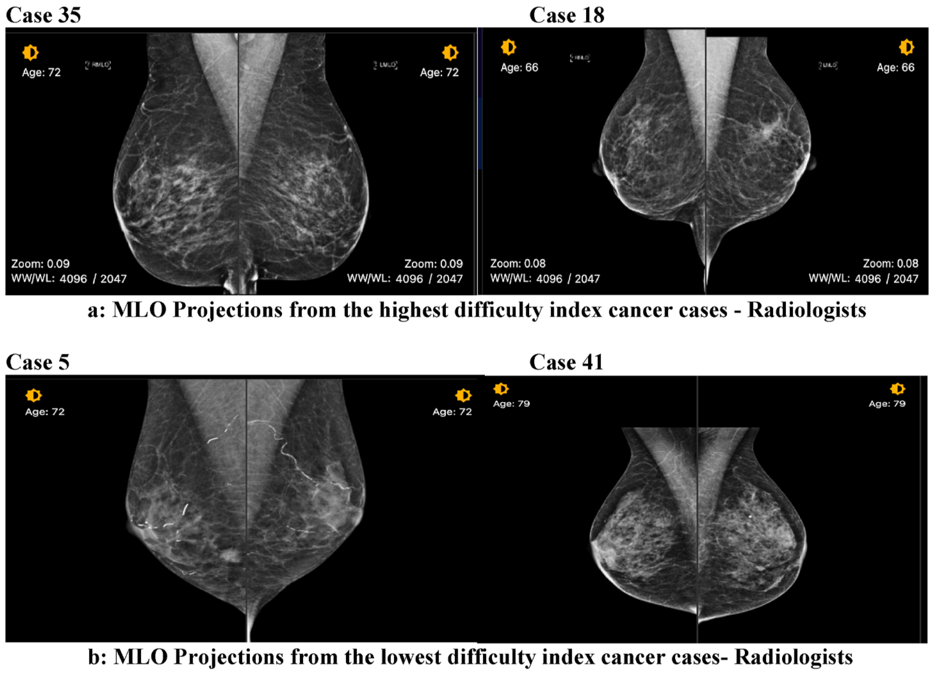

Figure 1. a MLO Projections from the highest difficulty index cancer cases - Radiologists. b MLO Projections from the lowest difficulty index cancer cases- Radiologists.

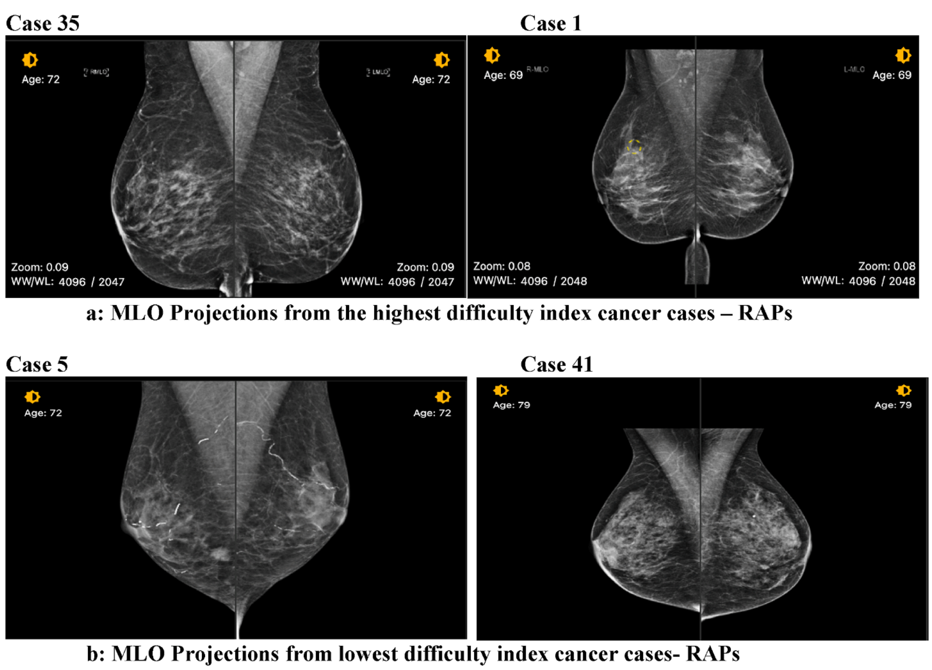

Figure 2. a MLO Projections from the highest difficulty index cancer cases – RAPs. b MLO Projections from lowest difficulty index cancer cases- RAPs.

calcifications ($p = 0.003$) (Table 4). No significant differences were noted for the radiologists.

Normal cases

For the normal cases, the four image tests sets with the highest and lowest difficulty indices are shown respectively for radiologists in Fig. 3a and b and 4a and 4b for the RAPs. Case 28 was identified by both professional groups to be the most difficult. Table 5 shows significantly greater difficult scores were shown for the RAPs for cases without prior images ($p = 0.03$)) compared to those with prior images respectively. No significant differences were noted for the radiologists.

Reader cohort comparison

The Spearman analysis demonstrated a strong difficulty indices agreement for cancer and normal cases respectively ($r = 0.83$ CI:0.61–0.93) AND ($R = 0.73$; CI:0.54–0.85) between the radiologist and radiographer groups.

Table 4
Results from the statistical analysis for cancer cases. Scores for difficulty indices are shown as well as p values. Statistical difference is shown in bold.

Case type	Median	IQ range	p value
Radiologist			
Soft tissue vs	0.06	0.04–0.14	0.13
Calcification	0.03	0.01–0.08	
Low density vs	0.05	0.03–0.11	0.67
High density	0.08	0.03–0.14	
Prior vs	0.06	0.03–0.14	0.16
No prior	0.03	0.01–0.18	
Radiography advanced practitioner			
Soft tissue vs	0.16	0.08–0.39	**0.003**
Calcification	0.06	0.04–0.08	
Low density vs	0.08	0.04–0.16	0.09
High density	0.18	0.08–0.29	
Prior vs	0.12	0.07–0.60	0.33
No prior	0.08	0.04–0.12	

Discussion

Criteria that impact upon image interpretation have been identified through platforms such as the Personal Performance in Mammographic Screening (PERFORMS) in the UK and Breast Screen Reader Assessment Strategy (BREAST) in Australia. Literature has also demonstrated an array of factors which determine high diagnostic efficacy.[15–21] However, these factors focus on all reader types grouped together or solely radiologists. This study investigated two types of health care professionals that interpret mammograms with a specific focus on the mammographic feature type that presented most difficulty. Our findings demonstrate that whilst a significant correlation was shown between RAPs and radiologists for case difficulty for both cancer and normal cases, some case-specific factors that significantly affected performance appeared to be unique to RAPs.

When images displayed a cancer, the cases that appeared to challenge radiographer readers to a greater extent, were those that featured soft tissues lesions rather than micro-calcifications and those in breasts with higher rather than lower density (although this latter finding was not significant – p = 0.09). The soft tissue vs micro-calcification finding is an interesting observation since with the improvements in digital technology and the availability of high performing post-processing tools, the visibility of soft tissue lesions, even in higher densities should have arguably improved. Nonetheless, it appears that perception of these lesion types is still the major problem,[22] suggesting that further technical advances around enhancing soft tissue lesion visibility is required particularly when this lesion appearance is juxtaposed with dense breast parenchyma. It is reassuring that RAPs appear to have fewer problems with micro-calcifications compared with other lesion-types, since characterisations of these lesions are important for differentiating between benign and malignant appearances.[23–25]

These data specifically describing radiography performance are preliminary and do not attempt to explain why detecting soft tissue lesions or indeed interpreting dense breast tissue may be more difficult compared with other lesions. Even so, this study suggests that radiologists may not be faced with the same challenges as their radiography colleagues for this appearance. Until the reasons for this inter-professional difference are identified, we can only surmise that extra support to optimise radiography performance could be enhanced through education and test set materials that focus on these appearances. The need for supplementary technical developments as suggested above have also been highlighted.

For normal cases the absence of prior images appeared to be a major agent. Both cohorts of readers found cases without prior images difficult to interpret compared to when priors were available, with this finding only statistically significant for RAPs. Previous research has demonstrated that prior images can improve cancer detection rates and can reduce the mammography screening recall rates.[14,26,27] The highly variable nature of breast structures even from normal mammographic appearance can make interpretation extremely difficult,[28,29] therefore, being able to compare current with previous images allows for a rigorous temporal assessment whilst reducing to some extent the confounding nature of normal nuances. The current findings are aligned with other studies that have demonstrated a strong

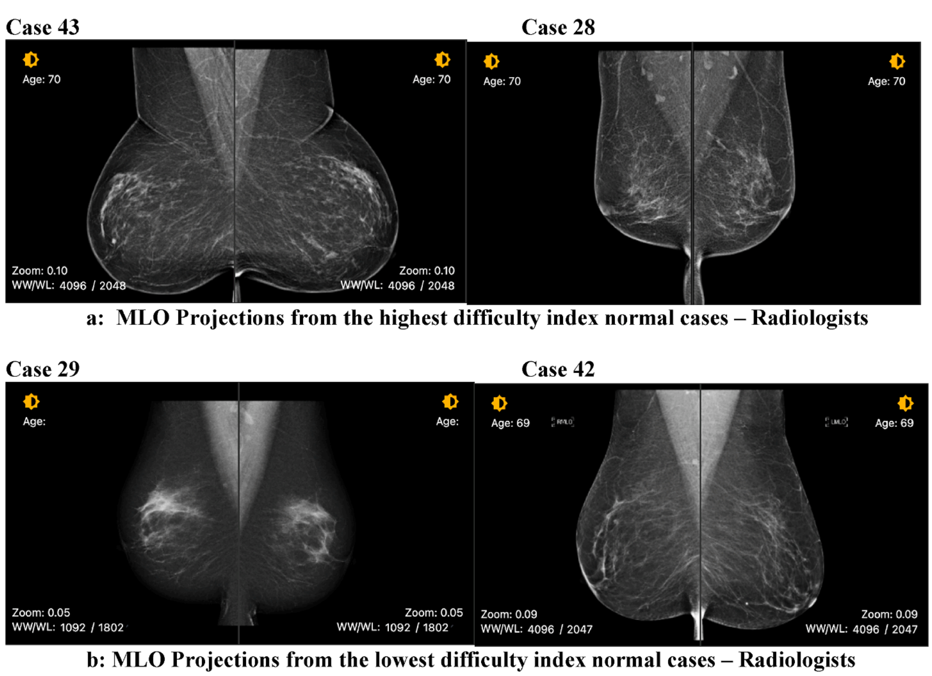

Figure 3. a MLO Projections from the highest difficulty index normal cases – Radiologists. b MLO Projections from the lowest difficulty index normal cases – Radiologists.

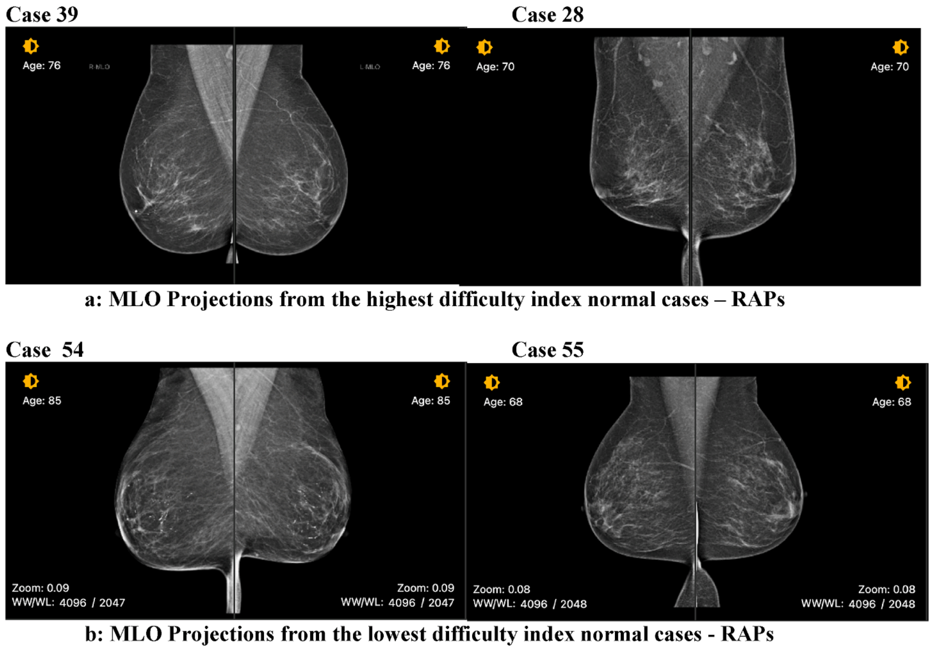

Figure 4. a MLO Projections from the highest difficulty index normal cases – RAPs. b MLO Projections from the lowest difficulty index normal cases - RAPs.

Table 5
Results from the statistical analysis for normal cases. Scores for difficulty indices are shown as well as p values. Statistical difference is shown in bold.

Case type	Mean/Median	SD/IQ range	p value
Radiologist			
Low density vs	0.17	0.05–0.30	0.97
High density	0.15	0.07–0.28	
Prior vs	0.15	0.04–0.26	0.20
No prior	0.27	0.09–0.38	
Radiography advanced practitioner			
Low density vs	0.20	0.09–0.32	0.50
High density	0.12	0.08–0.35	
Prior vs	0.12	0.08–0.32	**0.03**
No prior	0.24	0.15–0.45	

association between specificity and availability of previous mammograms.[26]

This is a preliminary study, and limitations include reader recruitment. Potential bias in performance needs to be acknowledged with variation in the number of participants in each reader group. The sample size of 16 RAPs and 25 radiologists may also be considered low, however similar studies have been reported with fewer participants.[30–32] Nevertheless, a larger group of both cohorts would have allowed for a more in-depth analysis. One way in which to increase our data is to work with test set groups in the UK and elsewhere such as PERFORMS (UK) and BREAST (Australia) respectively and establish if data available from previous readings there could complement our future work. This is now being planned. A recently established Special Interest Group affiliated with the Society and College of Radiographers and developed by the author (NL) will facilitate processes that will encourage larger radiography reader numbers in future studies.

Conclusion

In conclusion, this study demonstrates a reasonable correlation between radiologists and radiography advanced practitioners when determining which mammographic cases are difficult to interpret, regardless of whether the cases are with pathology or normal. This should help direct future AI efforts towards those cases that present with the highest difficulty ratings. Despite the agreement between the two professional groups, specific types of appearances were shown to be specifically challenging for the radiography advanced practitioners, which require tailored educational strategies and technical supports to minimise interpretative error.

Ethics approval and consent to participate

Ethical approval for this study was obtained from the University of Suffolk (RETH(P)21/006).

Written informed consent was obtained for anonymised participant information to be published in this article.

Availability of data

Data required for this study may be made available by the author(s) upon reasonable request.

Author contributions

NC: Conceptualisation, Data curation, Writing- Original Draft preparation, Investigation.
RS, CS, PB: Supervision.
MS, ZG: Software, Validation.

Declaration of Generative AI and AI-assisted technologies in the writing process

No AI tool was utilized.

Funding

No funding awarded.

Conflicts of interest

Noelle Clerkin is the niece of Professor Patrick Brennan.
Professor Patrick Brennan is CEO and Co-Founder of DetectedX.
Dr Moe Suleiman is Chief Technology Office and Co-Founder of DetectedX.

Acknowledgements

Not applicable.

References

1. Mirošević Š, Prins JB, Selič P, Zaletel Kragelj, Klemenc Ketiš Z. Prevalence and factors associated with unmet needs in post-treatment cancer survivors: A systematic review. *Eur J Cancer Care*. 2019;28(3). https://doi.org/10.1111/ecc.13060.
2. Elmore JG, Armstrong K, Lehman CD, Fletcher SW. Screening for breast cancer. *JAMA*. 2005;293(10):1245–1256. https://doi.org/10.1001/jama.293.10.1245.
3. Pane JF, Steiner ED, Baird MD, Hamilton LS, Pane JD. *Continued progress: promising evidence on personalized learning*. Rand Corporation; 2015. Available from: https://www.rand.org/pubs/research_reports/RR1365.html.
4. Geirhos R, Meding K, Wichmann FA. Beyond accuracy: quantifying trial-by-trial behaviour of CNNs and humans. *Neural Netw*. 2020;128:265–280. https://doi.org/10.1016/j.neunet.2020.05.002.
5. Panayides AS, Amini A, Filipovic ND, Sharma A, Tsaftaris SA, Young A, et al. AI in medical imaging informatics: current challenges and future directions. *IEEE Journal of Biomedical and Health Informatics [Internet]*. 2020 Jul 1;24(7): 1837–1857. Available from: https://ieeexplore.ieee.org/document/9103969.
6. Karadağ A, Hisar F, Göçmen Baykara Z, Çalışkan N, Karabulut H, Öztürk D. A longitudınal study on the effect of tailored training and counseling on the professional attitude of nursing students. *J Prof Nurs*. 2015 May;31(3): 262–270.
7. Tigchelaar A, Brouwer N, Vermunt JD. Tailor-made: towards a pedagogy for educating second-career teachers. *Educ Res Rev [Internet]*. 2010 Jan 1;5(2): 164–183. Available from: https://www.sciencedirect.com/science/article/pii/S1747938X09000487.
8. Hecksteden A, Meyer T. Personalized sports medicine – principles and tailored implementations in preventive and competitive sports. *Dtsch Z Sportmed*. 2018 Mar 1;2018(3):73–80.
9. Silacci A, Taiar R, Caon M. Towards an AI-Based tailored training planning for road cyclists: a case study. *Appl Sci*. 2020 Dec 30;11(1):313.
10. Juniarni C, Sodikin MA, Akhyar A, Almujahid A, Asvio N. The importance of personalized learning: how to tailor education to the individual needs of students. *Education Studies and Teaching Journal (EDUTECH) [Internet]*. 2024 Apr 30;1(1):188–196. Available from: https://journal.ppipbr.com/index.php/edutech/article/view/274.
11. Salman O, Khasawneh Y, Alqudah H, Alwaely S, Khasawneh M. Tailoring gamification to individual learners: a study on personalization variables for skill enhancement. *International Journal of Data and Network Science [Internet]*. 2024;8(2):789–796 [cited 2024 Sep 15] Available from: https://growingscience.com/beta/ijds/6664-tailoring-gamification-to-individual-learners-a-study-on-personalization-variables-for-skill-enhancement.html.
12. Gandomkar Z, Suleiman ME, Rickard MT, Jones C, Brennan P. *AI-driven personalised mammography education Cambridge conference on breast cancer imaging* 2024.
13. Taylor C, Monga N, Johnson C, Hawley JR, Patel M. Artificial intelligence applications in breast imaging: current status and future directions. *Diagnostics*. 2023 Jun 13;13(12):2041, 1.
14. Rawashdeh MA, Lee WB, Bourne RM, Ryan EA, Pietrzyk MW, Reed WM, et al. Markers of good performance in mammography depend on number of annual readings. *Radiology*. 2013 Oct;269(1):61–67.
15. Trieu PD, Lewis SJ, Li T, Ho K, Tapia KA, Brennan PC. Reader characteristics and mammogram features associated with breast imaging reporting scores. *Br J Radiol*. 2020 Oct 1;93(1114):20200363.
16. Maram Alakhras, Al-Mousa DS, Brennan PJ, Amir Tavakoli Taba, Rula Abu-Taimai, Rickard M. Impact of digital breast tomosynthesis on readers with different experience. SAVE Proc. SPIE 11513, 15th.
17. Hoff SR, Myklebust TÅ, Lee CI, Hofvind S. Influence of mammography volume on radiologists' performance: results from BreastScreen Norway. *Radiology*. 2019 Aug;292(2):289–296.
18. Gandomkar Z, Ekpo EU, Chen Z, Suleiman ME, Li T, Tapia K, et al. Expert radiologist performance does not appear to impact upon their capability in perceiving the gist of the abnormal on mammograms. *Proc SPIE 15th International Workshop on Breast Imaging (IWBI2020)*. 2020 May 22;11513: 156–162.
19. Ganesan A, Alakhras M, Brennan PC, Mello-Thoms C. A review of factors influencing radiologists' visual search behaviour. *Journal of Medical Imaging and Radiation Oncology*. 2018 Sep 10;62(6):747–757.
20. Ganeshan D, Rosenkrantz AB, Bassett RL, Williams L, Lenchik L, Yang W. Burnout in academic radiologists in the United States. *Acad Radiol*. 2020 Sep;27(9):1274–1281.
21. Alshabibi AS, Suleiman ME, Tapia KA, Heard R, Brennan PC. Impact of time of day on radiology image interpretations. *Clin Radiol*. 2020 Oct;75(10): 746–756.
22. Peters G, Jones CM, Daniels K. Why is microcalcification missed on mammography? *Journal of Medical Imaging and Radiation Oncology*. 2012 Dec 28;57(1): 32–37.
23. Michell MJ. Current challenges in mammography screening and diagnostic assessment. *Diseases of the Brain, Head and Neck, Spine 2016-2019*. 2016: 343–345.
24. Wilkinson L, Thomas V, Sharma N. Microcalcification on mammography: approaches to interpretation and biopsy. *Br J Radiol [Internet]*. 2016;90(1069). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5605030/.
25. Hovda T, Tsuruda K, Hoff SR, Sahlberg KK, Hofvind S. Radiological review of prior screening mammograms of screen-detected breast cancer. *Eur Radiol*. 2020 Apr;31(4):2568–2579.
26. Hayward JH, Ray KM, Wisner DJ, Kornak J, Lin W, Joe BN, et al. Improving screening mammography outcomes through comparison with multiple prior mammograms. *Am J Roentgenol*. 2016 Oct;207(4):918–924.
27. Roelofs AAJ, Karssemeijer N, Wedekind N, Beck C, van Woudenberg S, Snoeren PR, et al. Importance of comparison of current and prior mammograms in breast cancer screening. *Radiology*. 2007 Jan;242(1):70–77.
28. Digabel-Chabay, Allioux C, Labbe-Devilliers C, Meingan P, Ricaud Couprie M. Distorsions architecturales et difficultés diagnostiques. *J Radiol*. 2004 Dec 1;85 (12):2099–2106.
29. Matsubara T, Ichikawa T, Hara T, Fujita H, Kasai S, Endo T, et al. Novel method for detecting mammographic architectural distortion based on concentration of mammary gland. *Int Congr Ser*. 2004 Jun;1268:867–871.
30. Moran S, Warren-Forward H. Can Australian radiographers assess screening mammograms accurately? First stage results from a four year prospective study. *Radiography*. 2016;22(2):e106–e111.
31. Nair A, Gartland N, Barton B, Jones D, Clements L, Screaton NJ, et al. Comparing the performance of trained radiographers against experienced radiologists in the UK lung cancer screening (UKLS) trial. *Br J Radiol*. 2016 Oct;89(1066): 20160301.
32. Woznitza N, Piper K, Burke S, Bothamley G. Chest X-ray interpretation by radiographers is not inferior to radiologists: a multireader, multicase comparison using AFROC (Jack-knife alternative free-response receiver operating characteristics) analysis. *Acad Radiol*. 2018 Dec 1;25(12): 1556–1563.